MCESA
Maricopa County Education Service Agency

Assessment Development Webcast Handout

To help you process all of the information provided in the Assessment Development webcast, consider these questions:

1. Who is MCESA?

2. Why have they facilitated the development of assessments for special area courses?

3. What are two things that support the quality of the assessments?

4. What part of the assessment development process contributes to valid content?

5. What part of the assessment development process contributes to fair and reliable items?

6. What effect could these assessments have on classroom practices?

7. What are you still wondering?

Suggested Follow-up Activities

- Refer to the FAQ section of the MCESA website for responses to common questions.
- Share what you have learned with colleagues.
- Work with a team to analyze item specifications and blueprints.

Webcast Links

http://mcesa.schoolwires.net//site/Default.aspx?PageID=190

References

Haladyna, T.M. (1997). *Writing test items to evaluate higher order thinking*. Boston: Allyn and Bacon.

Haladyna, T.M. (2004). *Developing and validating multiple-choice test items*. Mahwah, NJ: Lawrence Erlbaum Associates: Mahwah, NJ.

Rabinowitz, S., Sato, E., Berkes, E. (2011). *Choosing assessments for measuring student growth.* Report prepared for the Student Growth Measure Task Force. San Francisco: West Ed.

Possible answers to questions:

1. Maricopa County Education Service Agency, Office of the County Superintendent
2. To help districts implement ARS 15-203, To provide a valid and reliable test for Group B teachers
3. Advisement from research and assessment experts, Collaboration with a professional test writing company, Aligning assessments to standards, Following research-based guidelines for test writing
4. Alignment to standards, Content review by teachers, Field testing
5. Bias review, Content review, Field testing
6. Focused curriculum , Increased focus on vocabulary and content language, Information about student's knowledge, Clear expectations for students, Students feel more accountable

MCESA
Maricopa County Education Service Agency

Assessment Development
Webcast Script

Welcome. The purpose of this webcast is to familiarize you with the contexts and processes that contributed to the development of assessments used in your district for teachers of art, music, PE, band, choir, theater, and dance. These assessments are the result of a county-wide collaborative project coordinated by MCESA.

MCESA is the Maricopa County Education Service Agency. We are the office of the Maricopa County School Superintendent, Dr. Don Covey. MCESA has three areas of focus: Executive Leadership, which oversees school board elections and home schooling; Economic Management, which assists districts with financial management; and Educational Innovations, which services school districts and charter schools through grant funded initiatives and a regional training center. One of our many functions is to support school districts in the county with the implementation of state-wide initiatives and legislation.

In particular, the state of Arizona's legislature passed Arizona Revised Statute 15-203. This statute says…

> *"The State Board of Education shall on or before December 15, 2011 adopt and maintain a model framework for a teacher and principal evaluation instrument that includes quantitative data on student academic progress that accounts for between thirty-three percent and fifty percent of the evaluation outcomes and best practices for professional development and evaluator training. School districts and charter schools shall use an instrument that meets the data requirements established by the State Board of Education to annually evaluate individual teachers and principals beginning in school year 2012 – 2013."*

The *Arizona Framework for Measuring Educator Effectiveness* is a 37-page document describing the details and expectations of this statute. In this document teachers are categorized as either Group A teachers, those with valid and reliable classroom level student achievement data aligned to the state standards or as Group B, those with limited or no available valid and reliable classroom level data aligned to state standards.

In a survey conducted in the fall of 2011, MCESA found that 74% of Maricopa County districts requested assistance with the development of assessments to comply with ARS 15-203, and thus began the project to develop assessments to support teachers in Group B. Press pause here to discuss the first two questions on the handout that accompanies this webcast.

- Who is MCESA?
- Why have they facilitated the development of assessments for special area courses?

The very first step was to create the Cross District Assessment Advisory Council. This council included administrators for research and assessment from across county school districts and nationally known consultants in the field of assessment. This council helped to determine the design and format of the project and the assessments.

When designing assessments, it is necessary to consider the whole framework of balanced assessment as a way to guide decisions. Any balanced assessment plan recognizes the value of a variety of formats including selected response items such as multiple choice or true-false items, constructed response items such as essays and short answer and performance-based assessments. All of these formats were considered when initially planning the design of the assessments to be created. Multiple choice tests were chosen for the first stage of the development process for several reasons. They are cost effective to create. They can be administered in a secure on-line format. They require minimal training for test administration. They align to the format of present national and high stakes tests. They are easy to score, and they provide a good base for critical thinking. Also, the cost of training and implementing constructed response and performance-based assessments were prohibitive for many districts in our current economic environment.

Given all of these strong reasons, it was decided to begin our collective assessment work with multiple choice tests of approximately 45 questions to be administered in one class period with the intention to investigate the use of performance-based assessments in the future. Furthermore, MCESA partnered with West ED, an assessment development company, to design and employ a research-based process to create assessments that are highly reliable and valid to ensure assessments linked to teacher effectiveness are of the highest quality.

Reliability means the degree to which a test produces similar scores each time it is used. Validity means the extent to which a test measures what it is intended to measure. To create valid and reliable test items, MCESA and WestED employed three key strategies in the development process. 1.) Align assessments to grade and content standards and Depth of Knowledge levels. 2.) Use professional editors to monitor style and formatting. 3.) Follow research-based guidelines for writing stems, answers and distractors in collaboration with special area teachers. Press pause here to discuss the third question on the handout that accompanies this webcast.

- What are two things that support the quality of the assessments?

Let's look more closely at the development process for the MCESA post-assessments. West Ed provided training and facilitated the work of more than 200 teachers and administrators representing 40 Maricopa County school districts. These teaching professionals worked to create 11 different assessments in these areas:

- Art for 3rd grade, 8th grade and high school
- PE for 3rd grade, 8th grade and high school
- Music for 3rd grade
- Beginning band, beginning choir, high school theater, and high school dance

The first step was prioritizing standards. This required teams of teachers to sort and organize the Arizona State Standards for their content area. In this step, teams had to …

- Agree on the most important things students should know and be able to do,
- Agree on the most important things to assess, and
- Determine the relative instructional emphasis, which means to prioritize the standards.

Because the standards for art, dance, PE, band and choir are grouped into grade bands or proficiency levels, the teams had to spend considerable time discussing how the standards look at the different grades within the band. Once the standards were prioritized, teachers were able to use the prioritization to write an assessment blueprint. This document shows the priority of the strands and concepts of the standards through a percentage. For example this blueprint shows that 56% of the items, which is 25 out of 45 items, on the3rd grade music assessment will be testing Strand 1: Create. Further down on the blueprint you can see that all 25 of those questions will be connected to Concept 5 of Strand 1: Reading and notating music. The other concepts have a zero listed, because the test will not contain any items aligned to those concepts. An end-of-course or end-of-grade assessment would be unmanageable if it tested every performance objective. That is why teacher prioritization of the standards informed the blueprint. Also, not every performance objective in the special areas can be assessed using multiple choice. Thus, it is important to have a balanced assessment plan.

The next step was creating item specifications. Item specifications serve as a bridge between the standards and the actual test items.

- They provide interpretations and clarification of the standards;
- They define the assessable content for each standard; and
- They serve as a collection point for decision rules related to the standards.

To write an item specification, a team may have to either narrow or expand the content of the standard. Furthermore, they specify the essence of the level of cognition for each standard. For example this item spec from 8th grade art shows the standard, the specifications, and the Depth of Knowledge, abbreviated as DOK.

- DOK 1 means the standard focuses on facts and simple recall.
- DOK 2 means the standard focuses on reasoning, such as comparing or applying factual knowledge.
- DOK 3 means the standard focuses on complex thinking and the application of skill.

After the item specifications were created, work began on writing the actual test items. Each item contains a stem and four answer choices. One of the choices is the correct answer, and the other three choices are called distractors. Each item is also coded with a DOK level. Writing items at the DOK 2 level was encouraged. The vocabulary of each item was monitored to be grade-level appropriate and aligned to the content standards. For the MCESA assessments, 165 items were created for each content area. It is necessary to make so many items, because not all of them pass the reviews and field testing process. The goal of the item-writing process was to finish with two high-quality forms of each assessment.

Once the items were written, they were edited by the West Ed content and assessment editors, using the Arizona Style Guide. The items were also formatted to be loaded in the electronic delivery system. Copyright permissions for any texts or graphics were also obtained in this stage.

After editing, the items were put through a bias review. Approximately 25 community members representing the diversity of Maricopa County examined each item to evaluate whether or not any item created an unfair advantage or disadvantage for any subgroup of students. Items were examined for stereotypes, insensitivity, gender imbalance, and offensiveness. A content review immediately followed the bias review. Teams of teachers who are experts in their content area reviewed each item for clarity, alignment to the standard, DOK level and answer choices. They ensured that the correct answer was listed and the other choices were plausible and not too obvious. After the reviews, the test items were revised and edited once more, and actual test forms were created for field tests.

In May 2012, we field tested the post-assessments for 3rd grade art, music and PE, 8th grade art, 8th grade PE and choir. Twelve districts in Maricopa County participated including the REIL grant partners: Alhambra, Gila Bend, Nadaburg, Tolleson, and Isaac as well as Cave Creek, Dysart, Laveen, Pendergast, Phoenix Elementary, Wilson, and Saddle Mountain. Approximately 12,000 students in over 437 classes participated. Psychometric analysis of the field test data provided reliability coefficients for each test form. Basic statistical practice expects test scores to have a reliability coefficient of .80 or higher. You can see that the results from field testing meet that expectation, indicating these tests have high reliability.

The last step in the assessment development process is to create the operational forms of tests. Any items that did not perform well in the field test were removed and two forms of each test were created for operational use using the quality items from the field-tested versions of the assessment. The following summative or post-assessments will be available for operational use in the 2012-13 school year: 3rd grade art, music and PE and 8th grade art and PE.

In May of 2013, MCESA, ATI, and WestEd will field test the following assessments: beginning band and choir, as well as high school art, PE, theater and dance.

Considering the rigorous and research-based process used to develop these post-assessments, you can feel confident that these assessments are a quality product fulfilling one component of a balanced assessment plan. Press pause here to discuss the questions 4 and 5 on the handout that accompanies this webcast.

- What part of the assessment development process contributes to valid content?
- What part of the assessment development process contributes to fair and reliable items?

In addition to the post assessments described here, 11 corresponding pre-assessments were developed in the summer of 2012. These pre-assessments will be field tested in October of 2012. To learn more about assessment development, consider consulting the reference materials provided on the handout.

Also, MCESA will be offering a one-day workshop for teachers or leaders of art, music and PE on October 2nd or 3rd, 2012. The purpose of the workshop is to closely examine the actual item specifications and blueprints and learn how to use them as a curriculum and instruction support for teaching of the content standards. Contact your district administrator if you are interested in joining the workshop on one of these days. To help you process all of the information provided here, answer the remaining questions on your handout.

Thank you for your interest in assessment development.

Notes

MCESA
Maricopa County Education Service Agency

Item Specification Samples

This is an example of an item specification that expands the standard.

Code	Content Statement	Item Specification
S1C1-1	Effectively employ age-appropriate fundamental movement skills in order to successfully participate in a variety of modified physical activities.	Focus on the movement: be able to identify, describe, and apply fundamental movement skills. Fundamental movement skills include, but are **not** limited to: Body Management skills/Non Locomotor skills: Bending Twisting Landing Stretching Static balancing Locomotor skills: Crawling Running Galloping Walking Hopping Skipping Dodging Rolling Climbing Dynamic balancing Object Control skills/Manipulatives: Throwing Catching Striking Bouncing Dribbling Kicking

This is an example that narrows or limits the standard.

S2C1-204	Discuss the roles of various art world experts (e.g., critics, art historians, curators, archeologists, conservators and others).	Limit items to the definition of art-world experts and examples vs. non-examples – see ArtLex. Use this definition of expert: someone who has a vast knowledge of a specific topic. Identify the following art-world experts and their basic job description: Art critic: critiques art work, **both** written and verbal. Art historian: preserves the history of works of art. Art curator: selects and arranges art for display. Archeologist: discovers artifacts. Conservator: preserves and treats art objects. Art collector: a person who acquires artworks for personal collection. This is a complete list of the art-world experts to be assessed in the summative assessment.

MCESA
Maricopa County Education Service Agency

Checklist for Item Specifications

- ☐ Type of Assessment: Multiple Choice, Constructed Response, Performance-Based Assessment
  - MC – Right Answer
  - CR – Multiple Interpretation/Perspective; Requires justification
  - PBA – References multiple resources; Completed over time; DOK 4

- ☐ Style: Learning statements begins with a verb

- ☐ Scope: Statements describe the domain or range of the PO, including sub-skills or concepts
  - Unpack all POs
  - Ensure that all learning statements are aligned to the intent of the PO. (How deep? How specific based on time available?)

- ☐ Critical Thinking:
  - Statements describe the cognition or critical thinking skills.
  - Include essence of DOK for the highest level of performance.

- ☐ Possible assessment question stems:

- ☐ Organization:
  - Connected PO's are written below the original PO with the label "Integrated POs."
  - Include Strand, Concept, and PO number with integrated PO. (Unwrap Integrated POs as well as original PO.)
  - Keep the highest cognitive demand learning statement/objective at the top. (PO or clarification of PO)
  - Sequence the statements below from simple to complex.
  - Include sub-skills in the learning statements where needed. (May be the PO)

MCESA
Maricopa County Education Service Agency

Blank Item Specifications Chart

Standard	Item Specifications

Depth of Knowledge

Level 1	Level 2	Level 3	Level 4
Level 1 involves recall and the response is automatic. Students either know the answer or not. Level 1 activities require students to demonstrate a rote response, follow a set of procedures, or perform simple calculations.	Level 2 activities require students to engage in mental processing and reasoning beyond a habitual response. These activities make students decide how to approach the problem, involving interpreting and developing relationships among concepts.	At Level 3 students are providing evidentiary support and reasoning for conclusions they draw. In most instances, having students explain and justify their thinking is at level 3. Level 3 activities have more than one correct response or approach to the problem.	Level 4 requires those tasks in which students must demonstrate reasoning, planning and developing connections within and beyond a content area. These activities usually occur over an extended period of time.
– Recall elements and details of story structure, such as sequence of events, character, plot and setting. – Conduct basic mathematical calculations. – Label locations on a map. – Represent in words or diagrams a scientific concept or relationship. – Perform routine procedures like measuring length or using punctuation marks correctly. – Describe the features of a place or people. – Demonstrate fingering of an instrument.	– Identify and summarize the major events in a narrative. – Use context cues to identify the meaning of unfamiliar words. – Solve routine multiple-step problems. – Describe the cause/effect of a particular event. – Identify patterns in events or behavior. – Formulate a routine problem given data and conditions. – Organize, represent and interpret data. – Play an instrument.	– Support ideas with details and examples. – Use voice appropriate to the purpose and audience in writing. – Identify research questions and design investigations for a scientific problem. – Develop a scientific model for a complex situation. – Determine the author's purpose and describe how it affects the interpretation of a reading selection. – Apply a concept in other contexts. – Compose melodies. – Plan art projects.	– Conduct a project that requires specifying a problem, designing and conducting an experiment, analyzing its data, and reporting results/solutions. – Apply a mathematical model to illuminate a problem or situation. – Analyze and synthesize information from multiple sources. – Describe and illustrate how common themes are found across texts from different cultures. – Design a mathematical model to inform and solve a practical or abstract situation. – Choreograph and perform a dance.

Courtesy of Southern Nevada Department of Professional Development and Webb, Norman L. and others. "Web Alignment Tool" 24 July 2005. Wisconsin Center of Educational Research. University of Wisconsin-Madison. 2 Feb. 2006. <http://www.wcer.wisc.edu/WAT/index.aspx>.

Reflection Collection

1. Write a one sentence summary defining the purpose and content of item specifications.

2. What ideas do you want to try for using the item specifications as a supportive curriculum tool?

3. How can you use the concept of Depth of Knowledge in your day-to-day planning?

4. Record 1-2 ideas for each category that you would like to implement at school in the next few months. (student actions, daily instruction, resources and supports)

MCESA
Maricopa County Education Service Agency

Here's What	So What	Now What?
• What did you notice about the color coding? • Did you have more green or pink? • How does your enacted curriculum compare with the written curriculum?	• What conclusion can you draw about the similarities and differences between your enacted curriculum and the written curriculum? • Looking at the assessment blueprint, how does the prioritization in the blueprint reinforce or vary from your usual practice? • How could the information on the blueprint impact your enacted curriculum?	• How can you use item specifications as a reference for planning or as a supportive curriculum tool?

MCESA
Maricopa County Education Service Agency

Samples from Observation Rubrics and Standards

Critical Thinking Elements

Source	Domain/Standard	Indicator/Evidence
InTASC Standards (Interstate Teacher Assessment and Support Consortium)	Instructional Strategies	Teacher engages all learners in developing higher order questioning skills and metacognitive processes.
InTASC Standards	Applications of Content	Teacher engages learners in generating and evaluating new ideas and novel approaches, seeking inventive solutions to problems and developing original work.
AZ K-12 Center Standards	Instructional Strategies	Integrates multiple levels of engagement and cognition methods. Higher order discourse and learning are frequently present.
REIL Learning Observation Instrument (LOI)	Critical Thinking	Student uses complex reasoning to make new meaning not provided by the teacher.
Marzano Teacher Evaluation Model	Effective Teaching Practices	Teacher asks questions or engages students in activities that require elaborative inferences; students use higher level thinking skills.
Stronge–based rubrics	Instructional Delivery	Engages and maintains students in active learning.
Maricopa County district A rubric	Engagement	Varying questions and activities promoting remembering, understanding, and applying with an emphasis towards analyzing, evaluating, or creating.
Maricopa County district B rubric	Critical Thinking	Contributes to students' conceptual understanding and promotes critical thinking, problem solving, and/or inquiry. Students make predictions, generate conjecture, or suggest alternative solution strategies.
Maricopa County district C rubric	Instruction: Activates Critical Thinking	Includes activities and instruction that support students in acquiring critical thinking skills at the application, analysis, synthesis, and evaluation level of Bloom's Taxonomy.
Marshall Rubric	Engagement	Gets all students highly involved in focused work in which they are active learners and problem solvers.

Sample Classroom Activities

DOK	Activity	Revisions
1	Students respond to this question chorally: What is the musical pattern heard in this piece of music?	
2	Turn and tell you neighbor what it means to dribble.	
2	Keep a personal dictionary of the new words we are learning all year.	
3	Students perform in an ensemble following a conductor's cues.	

MCESA
Maricopa County Education Service Agency

Workshop Evaluation

1. What questions do you still have?

2. What changes in your practice will you make?

3. What will you share with peers at your school or district?

4. What support can MCESA provide to you in the future?

5. What suggestions do you have for improving this workshop?